

# FDM darf nicht Weh tun – Ein Ansatz für eine den Forschungsprozess begleitende Repositoriumslösung

DINI/nestor AG Forschungsdaten-Workshop „Forschungsdatenrepositorien“  
am 27. und 28. November 2017 in der Universitätsbibliothek Stuttgart

AG Forschungsdatenmanagement | IT-SERVICES  
Ruhr-Universität Bochum, 44780 Bochum

Johannes Frenzel



- FDM? "Wir wissen, wie wir forschen."
- Metadatenanreicherung:  
keine 10-15 % Zeit für noch mehr Verwaltung
- Daten-Hoheit: "Gebe Daten aus der Hand"
- "Viel Arbeit Daten zu erzeugen"  
-> weitere eigene Auswertung/Analysen/Interpretation

Aber, (leider) auch ...

- Datenverlust (Total-/Teilverlust)
- Nachnutzbarkeit in Folgeprojekten ("Durchlauferhitzer")
- ...

## AG FDM

- Rektoratsprojekt: Universitätsbibliothek und IT.SERVICES
- Mission: Aufbau einer zentral koordinierten Forschungsdateninfrastruktur für die RUB
- Ressourcen: Personal (3, ab '18 +1 + 2xN.N.)

## Vernetzung und Kooperationen

- NFDI Bedarfe - Kooperationspartner bei UNEKE
- DMP - RDMO Software (UA Ruhr)
- (FDM)-Speicher (Backup Konsortium RWTH)
- Übergangslösung FDM-Speicher (RWTH u. UARuhr)

## Sechs UseCases - Forschungsspektrum RUB

- Einzelprojekte
- Forscherverbände (fach- und einrichtungsübergreifend)
- Verschiedene Projektstadien
- ca. 500 TiB in vier Jahren

## DMP Umfrage vor Projektstart SFB 1280

- Gemeinsame Fragen von SFB und AG-FDM
  - Vernetzung / Kooperation
  - Rechtliche Anforderung (Patientendaten)
  - Geräte, Methoden
  - Speicher, Software, Formate
  - ...

Filesystem (Samba-Share) + Backup (extern)

## Anforderungen

- Ordnerstruktur (wie Experiment, Simulation)
- "Metadaten" README-Datei (Minimalkonsens)
- Austausch zwischen ausgewählten Projekten
- Zugriffsbeschränkung (Rechte- und Rollen)
- kontrollierte Vokabulare (Suche, Auswertung, Meta-Studien)

? Repositoryum

"Heisse Daten" = Forschungsdaten !

- "Gute wissenschaftliche Praxis": Grundausrüstung (DFG)

Abholen/Begleiten des Forschungsprozesses

- Herausforderung: Heterogenität
  - Digitale Werkzeuge
  - Work-flows
- Übernahme von "Verwaltungsaufgaben"
  - Versionierung, Fixität, Backup
- 
- Vermeidung von Parallelstrukuten (grosse Daten)

## Kriterien an Repository/Front-End

- etabliert, "state-of-the-art", neue Technologie
- Schlüsselfertig, erprobt,
  - Personal
- lokale Lösung - Forschende entscheiden (RUB Policy)
  - Eigene Instanz Forschende/Gruppen (multi-tenancy)
  - Freigabe / Teilen - Forschende entscheiden
  - Rechte- und Rollenmanagement
- Datenmodell (Interoperabilität)
- Metadatenstandards
  - min. Dublin Core
  - Anpassung Forschungsvorhaben

## Kriterien an Repositoryum/Front-End

- ...
- Work-flow Rollen, Barriere-frei, intuitive Bedienung
- Schnittstellen
  - Hochschulbibliographie (OpenAIRE)
  - ORCID, DOI
- Import/Export (Exit-Strategie)
  - Migration -> NFDI?
- Community Software
  - Support, Anpassungen, Implementierung, Technologiewechsel

## Software

- Hyku (Samvera-Community, vormals Hydra)
  - basierend auf Hyrax (Curation Concerns, Sufia)
  - HykuDirect (gehosted von DuraSpace)
- ObjectStore, Cloud, IIF, Portland Common Data Model (PCDM) by FEDORA, Content-type models based on PCDM

## Status Tests

- Hyku
  - Daten einer FG Teilprojekts Zeitschriften/Bild-Scanns
- Speicher Interims-Lösung: RUB - UARuhr
- Anbindung Fedora

- (Beschaffung Storage-Hardware)
- Produktionstests
- Hands-On Testlauf mit UseCase
- (Re-)Evaluierung (Kooperation mit UNEKE)
- Weitere UseCases

## Langfristig ...

- Automatisierung, z.B.  
"machine actionable" DMP - RDMO Part 2 ?
- "Organisches Angebot" FDM:  
Immer kürze Zyklen: Software, Methoden, Experimente, Hardware
- ...